·科学技术文化研究·

人工“智”能的智慧、机智与明智

The Wit, Tact and Wisdom of AI

陈凡 /CHEN Fan[1]　吴怡 /WU Yi[2]

（东北大学科学技术哲学研究中心，辽宁沈阳，110169）
(School of Marxism, Northeastern University, Shenyang, Liaoning, 110169)

摘　要：通过对“智能”一词进行共时性和历时性的分析，可以将人工智能中的“智”解释为智慧、机智与明智这三种衡量标准。在对人工智能“三智”的阐述中找到人类解放脑力劳动和超越自身限定的根据，试图从价值维度上澄清人们在人工智能中对于“智”的真正期待；并强调智慧先于明智会带来伦理风险、机智先于明智会导致短视决策，即从实践意义层面上论证“明智”才应是人工智能首要达到的“智”，进而把道德因素前置到其他智能因素上来，以此对智能的概念进行再界定且提供发展进路的方法论指导。

关键词：人工智能　智慧　机智　明智

Abstract: It is concluded that the “intelligence” of artificial intelligence may be interpreted as three standards: wit, tact and wisdom through the synchronic and diachronic analysis of the word “intelligence”. This paper attempts to clarify the conception of the “intelligence” of artificial intelligence from the value dimension through the elaboration of “Three kinds of intelligence” of artificial intelligence. It is demonstrated that “wisdom” should be the first goal of “intelligence” of artificial intelligence in practical dimension. We should put the moral factor in a more important position than other intelligence factors by redefine the concept of intelligence and provide the methodology guidance for the development of artificial intelligence.

Key Words: AI; Wit; Tact; Wisdom

中图分类号：N0: TP18　**文献标识码：**A　DOI: 10.15994/j.1000–0763.2021.12.014

玛格丽特·博登（Margaret Boden）曾说，“人工智能不是关于计算机的研究，而是关于思维与行为中的智力的研究。”当前人工智能的发展能够在逻辑、识别、决策等具体领域达到人的“聪明才智”，但是目前学界对于人工智能中“智”定义的边界仍是模糊的。如果按照共时性的智慧和历时性的机智和明智两种解释的划分，它究竟能够在多大程度上算得上是智慧、机智和明智呢？智慧既包括逻辑思维能力，亦包括情感联想能力；机智要求智在短时间内的灵活性；而明智要求智在长时间内的远见性。从三种不同“智”的标准来审视人工智能的发展进度，以此反思人在人工智能这项新技术中所真正需要与期待的“智”究竟为何物。

一、智能判断：“三智”作为标准的可能性

基金项目：国家社会科学基金重大项目“新时代中国特色工匠精神的科学传承与创新发展研究”（项目编号：21ZDA020）。
收稿日期：2020年9月30日
作者简介：陈　凡（1954–）男，辽宁沈阳人，东北大学科学技术哲学研究中心教授，研究方向为技术哲学、科学技术与社会。Email: chenfan336@sina.com
吴　怡（1995–）女，辽宁沈阳人，东北大学科学技术哲学研究中心博士研究生，研究方向为技术哲学。Email: 17824941993@163.com

关于“智”的理解，源自人们在人工智能语境下对于“什么是智能”这一古老问题的探讨。有的学者认为机器能够模拟人的智能行为就是智能，而有的学者则认为智能要求机器本身就具有诸如意识、自我、思维等意蕴的。冯·诺依曼（John von Neumann）就秉持前一种看法，而在约翰·塞尔（John Rogers Searle）看来，智能是与其质料的性质相关的，所以他对于智能的理解属于后一种。[1] 总的来说，对于智能的研究大致分为这样三种模式：一是知识阈值理论，即认为智能取决于知识的数量及其一般变化；二是思维理论，即认为智能的核心在于思维；三是进化理论，即诸如行走力、感知力、繁衍力等用控制取代表征的理论。[2] 综合这三种理论，我们可见智能是可以理解为具有理论意义上的认知力和具有实践意义上的能动力。因此我们可以将“智能”看作是“智力”和“能力”的结合，那么只有智与行兼备才可算得上是人工智能。但需要强调的是，智具有优先性。智要优先于行，行始发自智，智来触发行；另一方面，智较行而言，更具困难性。智涉及精神层面的事物，所以智也是较行而言更不易实现的。当把“智能”一词从“人工智能”一词中脱离出来后，我们开始转向对于人类特性进行描述；继而又把“智”从“智能”一词中脱离出来，我们同时又暂且搁置了人类特性中的实践因素。实际上，在人工智能的语境下“智能”一词并不能被分割，因为当我们设计人工物之时就是希望它能够做出智能的行为。而我们为了更好地理解智能究竟是什么，可以暂时将重点放在智能中对于“智”的解释上：人的“智”有很多种，其中我们想要人工智能实现的无非涉及智慧、机智与明智这三种较为重要的智。其中，智慧属于从共时性的角度作为智能的衡量标准，而机智和明智属于从历时性的角度作为智能的衡量标准。

1. 智慧的解析

人工智能在物理符号主义之后走向联结主义是必然的，因为智慧本就是专属于人的，而想要机器具有智慧，对人的智慧进行模拟便是必不可少的。智慧首先要与智力一词区分开。智慧谓之“道”，而智力谓之“器”。智力只不过是人作为生命的部分技能，而人工智能需要机器去模拟人的智慧是指生物所具有的基于神经器官一种高级的综合能力，其中含有：感知、知识、记忆、理解、联想、情感、逻辑、辨别、计算、分析、判断、文化、中庸、包容、决定等多种能力。苏格拉底表明自己具有属人智慧，伯奈特将其解读为逻辑学、知识学说和伦理学的智慧；而儒家的中庸、道家的无为、墨家的兼爱等思想也皆是智慧。董军将智慧理解为人思维的综合，大成智慧学中将智慧分为量智与性智，其中量智主要指涉思维科学，而性智主要指涉文艺美学等。[3] 而钟义信将人类智慧分为显性智慧与隐性智慧两个部分，其中隐性智慧主要确定发现或创新的大方向，而显性智慧则是在这种方向之上进行具体的求解。当前人工智能可以根据给定的程序进行计算与分析，实际上就是在模拟人们的显性智慧，是在人类隐性智慧工作框架下的求解工具。[4] 也就是说，智慧是关涉人和世界的运维把握，它不仅包括理论的知识，也包括实践的知识；不仅关乎逻辑与识别，也含括了情感与联想。由此可以将智慧大致分为两个层面，第一层面包括理性能力，第二层面包括感性能力。人所具有的“智”不仅仅包括脑，还包括心。人的智慧是心与脑并用的智慧，固然既具有显性智慧，又具有隐性智慧。其中脑的部分大致可以理解为智慧的理性能力，而心的部分大致可以归结于智慧的感性能力。对于身—心关系，当代同一性理论拥趸认为心理状态与大脑相同。大脑神经化学研究也已经表明，大脑研究、精神病学与身—心理论之间有明确关联。这意味着智慧这两层意蕴并不总是孤立的，而经常是作为一体交叉出现的。所以这样看来，判断机器或系统是否智能可以根据智慧的理性能力和智慧的感性能力以及对于二者关系的运用把握程度来作为标准。

2. 机智的界定

机，巧也。机智具有“在复杂微妙的情境中迅速有把握和恰当行动的能力”，这意味着快速的准确决策，它也包含两层含义，一是快速做出决策，二是在快速的基础上正确决策。霍尔伯特·西蒙（Herbert Simon）和艾伦·纽威尔（Allen New–ell）都认为，“衡量一个系统的智能水平，是看它在面临任务环境所设置的种种变动、困难和复杂性时，达到规定目的的能力。”举个例子：中央处理器（CPU）和图形处理器（GPU）是两种不同性质的处理器。众所周知，CPU作为一般计算机系统的运算和控制核心，通用性很强。但相对CPU来

说，GPU的运行速度更快，其性能更加适用于专项工程。经过综合这两种处理器的不同优势而发展出的通用计算GPU（general purpose GPU），它虽然仅仅适用于一些特定的操作，但是可以大大提高其运行速度，比如深度学习——卷积神经网络CNN就是深度学习的一个成功的应用典范。由此不难看出，判断机器或系统是否智能可以将其是否能够在短时间内具有灵活性和应变能力作为标准。

3. 明智的内涵

“悲積習之生常兮，固明智之所别”，明智常与“理智之德”或“德智”相关。[5] 里夫（C. D. C. Reeve）认为“明智处于一个复杂的概念网之中，它与诸多概念有关，诸如感觉、科学–知识、制作–知识、考虑、决断、道德德性、智慧,以及理解、聪明、自然意义上的德性等”[6] 古希腊哲学的四枢德之一就是明智。亚里士多德指出，明智是一种关于人类践行的知识，是一种专门的思维，是依理性而实践的品质，是关乎道德的实际事务中进行思量的能力，是在实践中对具体事物的审慎的思考与把握的能力。[7] 由此可见，明智既包含思维的力量，又涵盖实践的维度;既涉及对“真”的把握，又具有对“善”的指向。尤其注意的是，明智是包含道德因素的一种智慧。因为它的远见性表明了这种智慧必定是与短视的利益驱动相反的。有些道德行为出自功利计算，有些道德行为则是出自自然正义，但我们不难发现，无论是出于何种目的，富有远见的智慧都能最终得出道德的决策。所以，判断机器或系统是否智能可以将其是否具有长远性、或者说在长时间内具有有效性和道德的价值意义作为标准。

迄今为止，人工智能发展的最前沿领域可以在具体的专业领域达到高超的识别水平和实现精准的数据分析，但它仍处于弱人工智能阶段。如依赖于大数据的机器学习可以轻松学习数据所喂养的外界事物，却难以学习情感，算不上是完全的智慧；caffe深度学习框架可以在一天之内只用一个GPU处理六千多万个图像，但这种惊人的速度也并不算完全的机智；道德人工智能体的构想还存在着理想与现实的差距:从“操作型道德体”到“功能型道德体”（这类机器人有能力对与道德挑战相关的情况进行评估和做出反应）再到“完全的道德主体”是一个循序渐进的过程，[8] 与明智的程度相差甚远。由此可见，当前的人工智能发展已经能够达到智慧的第一层面和机智的第一层面，但是还不能完美的兼具二者的第二层面，并且没有做到明智。

二、智能期待：“三智”作为目的的价值意义

西方哲学以二元对立的知识论看待技术价值，导致了技术乐观主义与技术悲观主义的对立。而马克思主义哲学从人的现实历史存在出发，认为技术价值不仅体现了满足人的需要，还体现了人的主动追求。它将技术视为一种文化形式，认为技术价值基本上是由其文化价值来决定的。文化价值意味着人对自己生命存在的文化意义的理解和认同。[9] 在人类中心主义看来,一切都是为“人”服务的。对此，人工智能的价值何在？或者换个说法，人究竟需要人工智能来做些什么？人们本可以在没有人工智能的时代就一直这样生活下去，但是人们为了生活得更好、为了实现人的自由和解放而选择创造这样一种技术作为手段。这种手段之所以称之为手段，是因为它不本身不是目的。人并不希望建造出与自己抢夺资源的另一个自己，人们甚至恐惧造出相似的自己、担心造出比自己更加优秀的自己后人的地位被动摇。西蒙指出，人工智能属于实验科学的范畴，研究人工智能的最终目的在于应用。[10] 所以，开发人工智能目的只可能在于将其应用于人类自身：一方面，人工智能有望为解放人的脑力劳动节省出时间，从而实现人们的自由愿望；另一方面，人工智能有望为人类生活注入新生力量，体现出人类向往的发展走向。

恩斯特·卡普（Ernst Kapp）的“器官投影说”表明，技术与人体之间具有密切的联系，人是制造器物的尺度。以往传统技术只是作为人类肢体延展的工具，解放了人们重复的简单体力劳动。而人工智能的开发正是从人的心与脑出发或者说是从人的整体智能因素出发形成的技术，试图解放人的脑力劳动。它是历史发展的必然结果，也昭示着人类发展的必然走向。根据马克思的自由全面发展理论，人达成自由全面发展这一最终目的是要通过自由时间的增多来实现的。自由时间的增多就意味着必要劳动时间的减少。以往传统

技术为人们解放体力劳动做出贡献，而人工智能为解放人们脑力劳动提供可能。若想依靠人工智能来解放人的脑力劳动，就须得让人工智能拥有人脑的智慧、机智与明智。看似轻松简单的人类脑力劳动的背后，实则隐藏着人作为类存在进化而来的智的客观能力与智的主观需要的吻合。当我们想要机器拥有智能从而符合人类对于智的成果需要之时，既需要对于人的智慧、机智与明智进行模拟，同时又需要明确人对于智慧、机智与明智价值的需要。

其一，人们寻求人工智能来解放自身的脑力劳动，势必会发展到进而解放自身的智力劳动，这意味着人工智能须得满足人们对于智慧的需要。可以说，能够满足人的智慧需要的机器一定是通用型的人工智能。当前人工智能之所以被洛夫莱斯夫人所诟病，就是因为它不可以实现创造。而我们并不只是需要一个在给定框架内进行具体操作的机器，我们仍寄希望于人工智能机器可以实现“用心”地为人类服务。可见，人对于智慧的需要不仅体现在“深蓝”所擅长的数学处理抑或是“沃森”所擅长的语言常识等那种理性智慧能力上，也体现在诸如联想创造、直觉审美等方面的感性智慧能力上。

其二，如今依赖于大数据而发展的人工智能终究摆脱不了经验论的局限，而人们需要的是在一个动态系统中面对各种具体应用时都能够灵活应对的机器，这就意味着人工智能还须得满足人们对于机智的需要。董军认为人的某些直觉判断是通过一些经验知识获得的。那么为什么这种经验在人应用时可以做到举一反三、融会贯通，而机器却时常不能将有限的经验知识转化为应对各种其他未曾出现过的具体应用的密钥呢？无法在新境遇下灵活应对问题就无法真正承担起解放人类脑力劳动和智力劳动的重任，所以单单是这一点，人们对于人工智能的机智需要就尤为凸显了。

其三，当前人工智能的发展既存在着物理伤害的风险又存在着伦理道德的风险，能否用长远的眼光考量问题并且做出明智的决策是人们最为关注的话题，这便意味着人工智能须得满足人们对于明智的需要。在一篇名为“信任AI需要什么？”的访谈中，马克·米切尔（Mark Micire）提到在消防和救援等方面对于一般技术设备的信任可以归结为可靠性和无害的能力。可靠性是指设备随着时间的推移重复执行相同任务并获得一致而准确的结果的能力，而无害的能力意味着它的安全性和低风险性。作为具有特殊性的人工智能与一般技术相比，除了可靠性与无害性以外，设备的透明性和可解释性对于人的信任来讲也是非常重要的。例如，在下围棋、玩游戏等具体领域方面，人工智能可以不必具有多么高超的“智”便可以轻轻松松地做到，然而在它给出人类不能完全断定的决策情况下，其不透明的决策过程使得人们难以信任。米切尔认为，即便是设备做出了错误的决策，只要我们能够解释其缘由，也是可以对于这种错误引发的信任危机进行修复的。[11] 可见，人们对于人工智能的信任需要可以大致分为长时间的可靠性、旨向“善”的无害性与可理解的透明性这三个重要的部分。明智就是这样一种可经得起时间与实践的考验、可应用且适用于人类的智，可见人们对于人工智能中明智的需要与期待是不言而喻的。

不仅如此，人总是有限与无限的统一。马克思所指的技术价值不仅在于满足人的需要，还体现了人的主动追求，其中的主动追求正是对人的超越性本质的描述。人们构建出另一个更加完满的“自己”，定然是对这个设定有所期许的：根据对人的超越性本质分析，人工智能技术自身就承载着“无限、理想、超越”价值意义。其中，智慧、机智与明智就是达成这种价值意义的关键。

具体说来，一方面，人虽与其他动物相比，是有其未特定化的优势所在，理性与智慧就是其最明显的表现。但人的智慧也终究是有限的、并非是无所不能的，人的身体固有生物性的局限，所以会在其有限的身体结构之外想要寻求另一种方式的永恒和超越。人工智能在某些具体领域地发挥可以超过人类生物的局限，也可以在极致精准的数据下给出事物的可能性，这是人类在人工智能中渴望得到弥补和证明的。另一方面，一旦人工智能实现了短时间的灵活性以及长时间的有效性，增强了人们对于它的信任，它就能够在一定意义上做为道德主体进而与人进行平等地交互，并实现人类对于理想状态的追求。换言之，人工智能通过对于智慧、机智与明智的达成，即可进而达成人类对于自身的“超我”演绎，这种将超我之于技术的方式为自我理想和道德良心提供现实基础的可能。

所以不难看出，人工智能所展现的智慧、机智与明智作为替代人类脑力劳动的必要条件，可以缩减人们的必要劳动时间、为人们创造更多自由时间，其技术价值是积极可观的。而且，在隐喻的视角下看来，具有智慧、机智与明智的人工智能体就好比新生命一般，凝入社会之时也重构了人们的文化背景。

三、智能实践：明智作为“三智”之首的必要性

智慧、机智和明智三者对于人工智能而言都是不可或缺的。但是究其根本，明智是最为关键的：其一，起初旨在模拟人类智慧的机器容易走向不可控的局面，所以从一开始就应让机器具有“明智”。其二，不具有明智的机器无法实现“机智”，随着硬件的跟进短时间内让机器迅速做出决策并不难，而难点在于确保机器在短时间内决策的正确性。

智慧先于明智会带来伦理风险。众所周知，现阶段人工智能发展到瓶颈期仍是由于身–心关系问题的不明晰性。所以大多数学者尝试着从脑科学、数学等方式对此进行研究来使得人工智能具有兼具理性能力与感性能力的“智慧”。但现在问题的关键并不在于此，因为我们不能先去设想一个有风险的事物的可能性再去分析其可行性，而应想去设想其是否具有可行性继而再去分析其实现的可能性。由于明智包含着道德因素,所以明智应该在设计“智慧”之前首要考虑的。前面提到，脑的部分主要承担智慧理性的职能,而智慧感性大多体现在“用心”。其中，如果智慧感性的用心能够具有善的指向，那么则是从义务论的角度实现从智慧到明智的过渡；如果智慧理性的“用脑”能够通过计算将道德的因素考虑进去，那么则是从功利主义的角度实现从智慧到明智的过渡。无论是从功利主义出发，还是从义务论出发，实际上都是基于人们对于人工智能的信任需要这一无害性前提。雷蒙德·库兹韦尔的奇点理论已经使我们意识到人工智能发展速度呈指数增长，它表明人工智能一旦具有“智慧”的那刻起便容易走向人类不可控的局面。这时的办法就只有让人工智能自己控制自己，所以我们应将研究的重点从“智慧”转向“明智”。

机智先于明智会造成短视决策。机智的决策说明在当下的决策是迅速而准确的。有时候我们会需要机器在短时间内给出较为精准的答案：比如在刑侦方面，可以利用大数据和人工智能来进行犯罪人识别、犯罪行为监控、犯罪情报分析等来缩短破案时间和分析成本。除此之外，还有很多决策都是需要人工智能的机智来完成。诸如意念购物眼镜，当你带上有传感器的眼镜，人工智能会记录看到某物品时的人体感应数据，计算当前感应器数据的对应的购物欲分类概率。如果检测到你对刚看到物品的购买欲超过80%，它将自动地把商品信息加入购物车列表，实现消费领域的机智；又如无人驾驶，相比之下它比人工驾驶更加精准和专注，能够借助光探测、测距传感器、近程碰撞传感器等组件，实时地收集丰富的交通环境数据，并通过智能算法有效减少交通事故和缓解城市拥堵，[12] 实现交通领域的机智。问题在于，消费领域的机智是否与消费者的隐私产生矛盾？交通领域的机智是否存在着高风险的安全隐患？很显然，这些人工智能所带来的机智恩惠很难回应其带来的相关伦理问题，而唯有将包含道德因素的“明智”设计置于“机智”设计之前，才能给“机智”确立一个人本的保障。

总而言之，美德与知识二者的关系不仅仅存在于人类自身的哲思中，也在人工智能这一技术领域里有所影射。我们常说一个文化水平高的坏人相比一个文化水平低的好人要更加可怕，同样一个具有高智却坏的机器要比一个低智却好的机器更有破坏性。所以，当我们将智慧、机智和明智这三种标准本身进行审视之时，我们不难发现明智的远见性和道德性使得其他智能标准都黯然失色。唯有将明智设计于其他智能因素之前，才可稳妥地实现人对于人工智能中“智”的真正需要。

明智先于智慧与机智不仅具有必要性，还具有一定的可能性。由于明智必然包含道德因素，而乔安娜·布赖森（Joanna Bryson）又认为，我们的道德植根于我们的生物学和我们的文化。[13] 也就是说，这种明智是具有先天的身体结构和后天的文化渲染两种作用下结合而成的。这便意味着，仅仅依靠联结主义的神经网络或是仅仅依靠行为主义的强化学习都是不能轻易达到这种明智的，而只有两者的巧妙结合或许才能为明智开辟出一条道路。鉴于这种实现进路上的可能性，也意味着人工智能中的明智意蕴是可以被前置在其他智的因素之上的。

结　论

通过对于人工智能中“智”的概念的分析，可以重新理解智能的内涵，并且澄清人们在这项新技术中的真正期待。第一，将智慧作为共时性的标准、将机智与明智作为历时性的标准来衡量人工智能的发展进度，能够结合定性分析与定量分析，将还原论与整体论相结合，使较弱人工智能和强人工智能的说法更加清楚和有迹可循。第二，智慧、机智与明智三种不同的智可以在价值维度上体现出人的特定需要，进一步明确了人工智能应然的技术价值以及它与人之间的共在关系。第三，将明智的地位提升至其他智能因素之上，具有伦理道德意义上的合理性。按照先明智后智慧、先明智后机智的研究顺序可以为人工智能的良性发展提供新思路。

总的来说，只有首先澄清人们对于“智”的真正需要与期待，才能将人工智能技术作为一种自我实现与超越的手段。并且只有明晰了不同“智”的概念，才不被简单的“智能”一词混淆其真正含义。明智是属于人类的大智慧，或许我们想要进一步开发人工智能使其更好地为人服务又不受制于它，只有从自身的“智”中去寻找依据才是可能的办法。

［参 考 文 献］

[1] 郭斌. 塞尔“中文屋”思想实验的哲学意蕴辨析 [J]. 自然辩证法研究, 2005, 21（12）: 29–32.

[2] 赵泽林. 人工智能的基础哲学问题探秘 [M]. 广州: 世界图书出版广东有限公司, 2012, 04.

[3] 熊十力. 返本开新 熊十力文选 [M]. 高瑞泉编选, 上海: 上海远东出版社, 1997, 1.

[4] 钟义信. 人工智能: 信息技术的制高点——献给《中兴通讯技术》创刊20周年 [J]. 中兴通讯技术, 2015, 21(3): 1–3.

[5] 曹志、彭丽君. “明智”德性研究 [J]. 长沙铁道学院学报（社会科学版）, 2011, 12（4）: 1–3.

[6] Reeve, C. D. C. *Practices of Reason: Aristotle's Nicomachean Ethics* [M]. New York: Clarendon Press, 1992, 67.

[7] 亚里士多德. 尼可马克伦理学 [M]. 廖申白译注, 北京: 商务出版社, 2009.

[8] 王东浩. 人工智能体引发的道德冲突和困境初探 [J]. 伦理学研究, 2014, 70（2）: 68–73.

[9] 于春玲. 论马克思对西方哲学技术价值困境的超越——基于文化哲学的视角 [J]. 东北大学学报（社会科学版）, 2012, 14（2）: 101–106.

[10] 董军. 人工智能哲学 [M]. 北京: 科学出版社, 2011, 109.

[11] Ray Mond Ka-Man Sheh. 'What is Required to Trust AI?' [J]. *KI–Künstliche Intelligenz*, 2018, 32(4): 279–281.

[12] Goodall, N. J. *Machine Ethics and AutoMated Vehicles* [M]. Meyer, G., Beiker, S. (Eds.)Road Vehicle Automation Berlin: Springer International Publishing, 2014, 93–102.

[13] Bryson, J. J. 'Patiency is not a Virtue: The Design of Intelligent Systems and Systems of Ethics' [J]. *Ethics and Information Technology*, 2018, 20(1): 15–26.

［责任编辑　孟建伟］